

REVIEWER A

El manuscrito estudia el efecto que puede tener la práctica de invertir ítems al medir constructos psicológicos. Para ello evalúa mediante técnicas de análisis factorial confirmatorio diferentes modelos. La conclusión del estudio es que se constata la existencia de un efecto de método asociado a los ítems revertidos.

El texto usa de forma indiscriminada los términos “ítem invertido” e “ítem negativo”, cuando en realidad se refieren a dos tipos de ítems diferentes. Por ejemplo, en la primera página del manuscrito se encuentra la frase: “Así, la práctica de invertir ítems se llevaba hasta el límite de recomendar un equilibrio entre el número de ítems positivos, en el sentido del constructo a medir, y negativos, en la dirección opuesta.”. Un ítem negativo es un ítem que presenta una frase con una negación: hoy en día su uso está fuertemente desaconsejado en los manuales de construcción de tests ya que son ítems confusos y con frecuencia difíciles de interpretar. Si la negación aparece en la primera palabra del ítem (es decir, ítems que empiezan con las palabras “no” o “nunca”), la ambigüedad del ítem se acentúa. Un ejemplo de ítem negativo es el ítem 5 del cuestionario utilizado por los autores. Por contra, un ítem revertido es un ítem que presenta una frase o afirmación positiva pero en dirección opuesta a la dirección en que se pretende medir el rasgo latente. Un ejemplo de ítem revertido es el ítem 3 del cuestionario. El manuscrito va más allá y, no sólo utiliza los dos términos como si se tratase de ítems de formato equivalente, si no que engloba y manipula estos dos tipos de ítems a efectos prácticos como ítems indistintos. Desde mi punto de vista, esta manipulación limita los resultados y las conclusiones que se obtienen. Se debería haber utilizado un test que no incluya ítems negativos, sino tan sólo ítems revertidos.

En el texto se hace una mera (y hasta cierto punto despectiva) referencia a la razón por la que se aconseja la utilización de ítems revertidos: (página 4) “Esta recomendación se basaba en la pretendida omnipresencia del sesgo de aquiescencia”. Dado que el sesgo de respuesta debido a respuestas aquiescentes está reconocido como uno de los sesgos más importantes en cuestionarios de personalidad (junto con la deseabilidad social), el manuscrito debería presentar de forma más formal este sesgo, sus consecuencias y los diferentes métodos que se han propuesto en la literatura para controlar este sesgo. Nótese que este sesgo de respuesta no es tan trivial como el manuscrito da a entender: si una proporción substancial de participantes responde un test sin ítems revertidos respondiendo sistemáticamente 4 ó 5 (al azar) sin ni tan sólo leer los ítems, las propiedades psicométricas del test aparecerán artificialmente como muy aceptables (contundente unidimensional, altas correlaciones ítem total, alta fiabilidad, etc.). Desde este punto de vista, el control de este tipo de respuestas es pues importante.

Si bien se evalúan diferentes modelos, no se ha considerado en el estudio ningún modelo que tenga como objeto controlar la variabilidad debida a sesgo de respuesta aquiescente. Al no evaluar ningún modelo de este tipo, el lector puede concluir que un modelo de dos factores ortogonales es más que razonable: un factor relacionado con la variable latente, y un factor relacionado con el sesgo de respuesta aquiescente presente en los ítems. Los trabajos de Billiet & McClendon son un ejemplo de cómo evaluar estos modelos en el contexto del análisis factorial confirmatorio.

El estudio sería interesante si demostrase que, después de controlar la variabilidad debida a sesgo de respuesta de aquiescencia, todavía aparecen efectos de método en ítems revertidos (que no negativos).

REVIEWER B

El presente trabajo de investigación aborda un tema que continúa siendo relevante hoy en día a pesar de la cantidad de estudios que lo han abordado previamente: ¿Son los factores compuestos por ítems inversos factores de rasgo o producto de artefactos metodológicos? En el caso actual la pregunta se trata de responder a partir de un análisis profundo mediante modelos SEM a las respuestas de la Escala de Autoestima de Rosenberg (RSE).

Como se comenta en el artículo, la escala RSE ha sido estudiada extensivamente en cuanto al tema de los ítems inversos y los factores de método se refiere, incluyendo múltiples investigaciones que han tratado de esclarecer la problemática a partir de matrices multirasgo-multimétodo (MTMM) modeladas mediante la técnica del análisis factorial confirmatorio (CFA). En este sentido, el trabajo actual ofrece como valor añadido el análisis de las MTMM a partir de las puntuaciones a dos versiones del RSE: la versión original del test y una versión compuesta sólo por ítems directos en la cual los ítems originales inversos han sido parafraseados y escritos en sentido positivo. Asimismo, cabe destacar que todos los participantes respondieron a ambas versiones de la escala.

A nivel general el estudio tiene el potencial de contribuir en el avance de la literatura acerca de ítems inversos y factores de método. No obstante, hay algunos puntos importantes que han de ser abordados para que éste sea capaz de llenar dicho potencial:

- 1) En la Introducción no queda claro cuál es el aporte del estudio o la razón que ha motivado su ejecución. Como bien se comenta en el manuscrito, la escala RSE ha sido estudiada en numerosas ocasiones para tratar de esclarecer la naturaleza del factor compuesto por sus ítems inversos. Entonces, ¿Cuál es el valor añadido del estudio actual? Si bien se menciona que se intenta “*clarificar el efecto de la formulación negativa de los ítems, utilizando dos versiones de la escala de autoestima de Rosenberg (1965), una en su formato original con ítems positivos y negativos (invertidos) y otra transformada de modo que todos los ítems sean positivos*”, no se explica el porqué al hacer esto se puede obtener una información diferente a la que ya se tiene a partir de los estudios que han utilizado los mismos modelos SEM pero que no han utilizado las dos versiones y con los mismos sujetos. En otras palabras, es importante señalar por qué este abordaje distinto puede esclarecer la problemática. En este sentido, habría que comentar el artículo de Greenberger, Chen, Dmitrieva y Farrugia (2003) titulado “*Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: do they matter?*” en el cual se aplicaron 3 versiones de la RSE (la original, una sólo con ítems en positivo y otra sólo con ítems en negativos) y señalar por qué el estudio actual extiende o aporta información nueva a los hallazgos del mismo.

2) A nivel metodológico hay un tema importante que no se comenta en el artículo y es que al analizar se manera conjunta las respuestas a las dos versiones del RSE se están identificando factores con ítems duplicados (los 5 ítems originales escritos en positivo). Se supone que esto ha de afectar de manera sustancial el ajuste de los modelos ya que son los mismos ítems aplicados en la misma batería, por lo cual sus errores estarán correlacionados. No obstante, en el artículo no se menciona si efectivamente se liberaron las covarianzas entre los errores de estos pares de ítems repetidos, y como no aparece la Figura 1 en el manuscrito (está citada pero no aparece al final) es imposible comprobar si se liberaron o no dichas covarianzas.

3) El modelo 3bis, que muestra el mejor ajuste de todos, surge *a posteriori* al liberar la covarianza entre los errores de la versión positiva y negativa del ítem 10. Ahora bien, surge la pregunta ¿Por qué no se ha liberado este parámetro en los otros modelos? Cabe pensar que los demás modelos, o por lo menos algunos, también mejorarían en ajuste de liberar este parámetro. En el manuscrito no se da una razón por la cual este parámetro ha sido liberado únicamente en el modelo 3.

4) Dado que se analizan muchos modelos parecería que la lectura sería más clara si se eliminaran los modelos 2, 5 y 6, y se analizaran únicamente los modelos 2bis, 5bis y 6bis. Los modelos 2, 5 y 6 estudian el ajuste de estructuras con factores de rasgo no correlacionados, cuando a nivel teórico todo sugiere que los factores de hecho lo estarán (y efectivamente la correlación estimada es mayor a 0.80). La lectura se haría más fluida si se eliminan estos modelos que más que aportar información le están añadiendo una complejidad innecesaria al manuscrito.

5) Al igual que como se comentó respecto a la Introducción, sería apropiado dedicar un espacio en la Discusión para comentar el porqué el trabajo actual constituye un valor añadido a la literatura sobre la RSE y el funcionamiento de los ítems inversos. Nuevamente, se comenta que se ha hecho algo diferente, pero no se dan las razones por las cuales esto constituye un aporte o permite responder a nuevas preguntas.

Luis Eduardo Garrido
Universidad Iberoamericana (República Dominicana)